Grid accounting service: state and future development

T Levshina[1,2], C Sehgal[2], B Bockelman[3], D Weitzel[3] and A Guru[3]
[2] Scientific Computing Division, Fermilab, Batavia, IL 60510, USA
[3] Holland Computer Center, University Nebraska – Lincoln, Lincoln, NE 68588, USA

E-mail: tlevshin@fnal.gov

Abstract. During the last decade, large-scale federated distributed infrastructures have been continually developed and expanded. One of the crucial components of a cyber-infrastructure is an accounting service that collects data related to resource utilization and identity of users using resources. The accounting service is important for verifying pledged resource allocation per particular groups and users, providing reports for funding agencies and resource providers, and understanding hardware provisioning requirements. It can also be used for end-to-end troubleshooting as well as billing purposes. In this work we describe Gratia, a federated accounting service jointly developed at Fermilab and Holland Computing Center at University of Nebraska-Lincoln. The Open Science Grid, Fermilab, HCC, and several other institutions have used Gratia in production for several years. The current development activities include expanding Virtual Machines provisioning information, XSEDE allocation usage accounting, and Campus Grids resource utilization. We also identify the direction of future work: improvement and expansion of Cloud accounting, persistent and elastic storage space allocation, and the incorporation of WAN and LAN network metrics.

1. Introduction

The Grid Accounting Service (Gratia) [1] is a federated accounting service that is used by the Open Science Grid, Fermilab, Holland Computing Center at University of Nebraska – Lincoln and other institutions to collect resource utilization records. The collected data includes information about batch and glide-in jobs wall duration, cores and cpu usage as well as user identity, submission and execution sites. Gratia also collects information about file transfers, storage allocations, and grid services availability. In this paper we will discuss the undergoing development as well as our future plans for Gratia extension.

2. Statistics and Current Deployment

The OSG production Gratia service has been deployed at Fermilab since 2007. It is operated by the Grid and Cloud Service Operation group. The Gratia service deployment is included in the FermiGrid High Availability infrastructure[2]. It provides a robust and scalable service that was able to sustain more than 200,000 records per hour rate for processing data. It has collected more than 1 billion individual usage records since 2005.

Gratia collects information by means of probes that are running on remote sites and reporting data to a collector. The list of probes includes:

[1] To whom any correspondence should be addressed.

- Batch (HTCondor, PBS, LSF, SGE, SLURM)
- Transfer (gridftp, hadoop, dCache, xrootd)
- Storage (dCache, xrootd, hadoop)
- Service Metrics
- Cloud Accounting
- Unix Accounting

Currently 130 sites are running Gratia probes and reporting resource usage.

3. Architecture Overview

The Gratia Service consists of several subsystems that include a collector, a reporter (WEB UI for user and administrators), and a database. The records gathered by various probes running on remote sites are sent to a collector by using the Gratia API. Gratia supports hierarchical collectors' structure and permits forwarding and filtering between collectors.

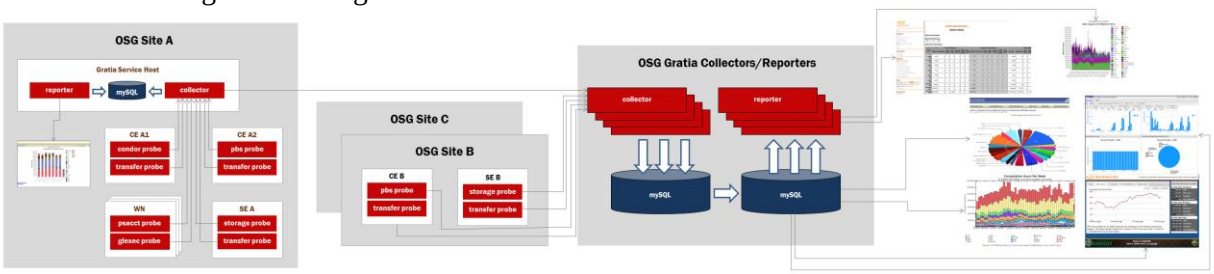

Figure 1: Gratia Architecture.

The Gratia framework provides means to generate various reports that allow graphical and textual access to the data. It also allows access to the data via several web portals.

The available metrics are used by users, VO and site administrators, OSG management, as well as funding agencies.

4. Cloud Accounting

The FermiCloud Service [3] is an infrastructure-as-a-service facility that provides the access to virtual machines on demand without a system administrator's intervention. It is used by developers, integrators and testers. It also allows starting on-demand services such as gridftp or mysql servers. We need to provide accounting information about cloud utilization. This information is collected by the Gratia cloud probe. We have currently implemented a probe that interfaces with Open Nebula 3.2. All information about the VM's owner, physical machine, states, start and stop times, memory and cpu provisioning could be acquired from interfaces provided by Open Nebula. The OGF Usage Record structure, used by Gratia, was sufficient to represent the VM accounting information. No changes in the Gratia collector code were required.

The cloud probe is running as a cron job and on start up requests information from Open Nebula using Open Nebula API. During the execution the information about the last known VM and a list of VMs that have changed their state is cached. This approach minimizes the load on the Open Nebula

server and allows faster generation of the cloud utilization records by a probe.

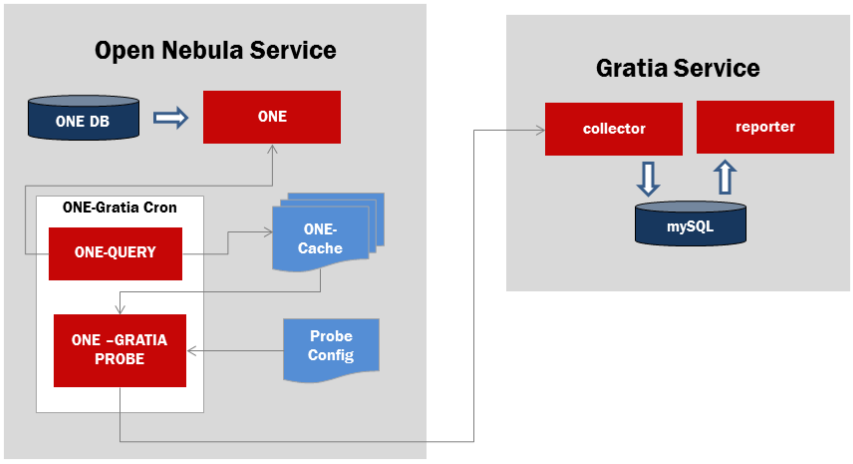

Figure 2: Gratia Cloud Probe Architecture.

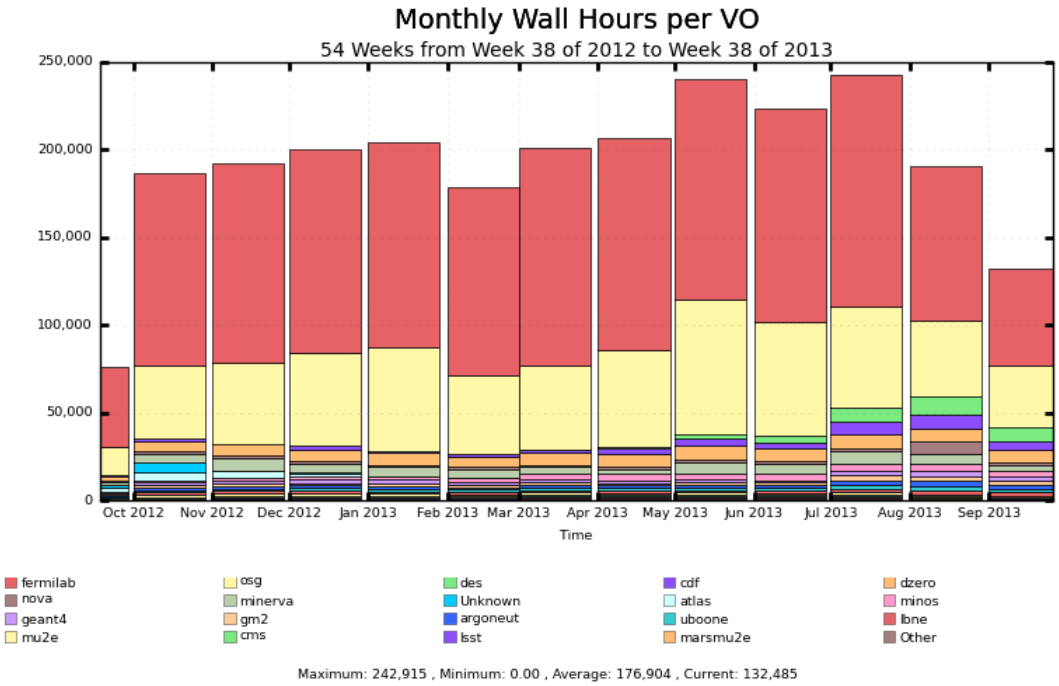

Figure 3: Monthly FermiCloud Resource Usage by various VOs.

5. Campus Accounting

The cyber-infrastructure is undergoing a self-reorganization to address the need for the increasing computational scale. Resources are reorganized in larger and independent conglomerates known as Campus Grids (CG). A Campus Infrastructure [4] provides a means to access shared resources. It enables researchers to transparently use various computational resources that include resources from other campuses as well as resources from the national cyber-infrastructure, such OSG or XSEDE[5]. The primary concern for Campus Grids accounting is the accurate recording of usage of the OSG

resources by CG users. We need to account for jobs that have been flocked to OSG resources from CG submitters.

A local campus user usually does not use a certificate. Several modifications were made to Gratia so probes can account for a local user's usage on campus and on OSG.We currently require that system administrators that maintain CG infrastructure register a new group of researchers in the OSG Information Management System (OIM) by providing minimal information about the nature of the research (Field of Science), the name of a Principal Investigator and the preferred Project Name. The project name is set by users during the submission of the job to the CG. The jobs records, submitted by CG users that contained specific Project Names are collected in Gratia. This way the resource usage can be traced back to the individual group.

Name	**SNOplus**
Description	
	SNO+ is a multi-purpose liquid scintillator detector with a primary goal of studying neutrino-less double beta decay in Tellurium-130, and is also capable of measurements involving solar neutrinos, reactor antineutrinos and geoneutrinos, supernovae, certain nucleon decay modes. Data collected by the detector are moved to (UK and Canadian) grid storage, where automated processing occurs. The large number of simulated data sets required for statistical analyses are also produced on grid resources.
Organization	**University of Pennsylvania**
Department	**Physics and AstronomySponsor**
Virtual Organization	**OSG**
Principal Investigator	**Joshua R Klein**
Field Of Science	**Physics - Neutrino**

Figure 4: An example of the project registration in OIM.

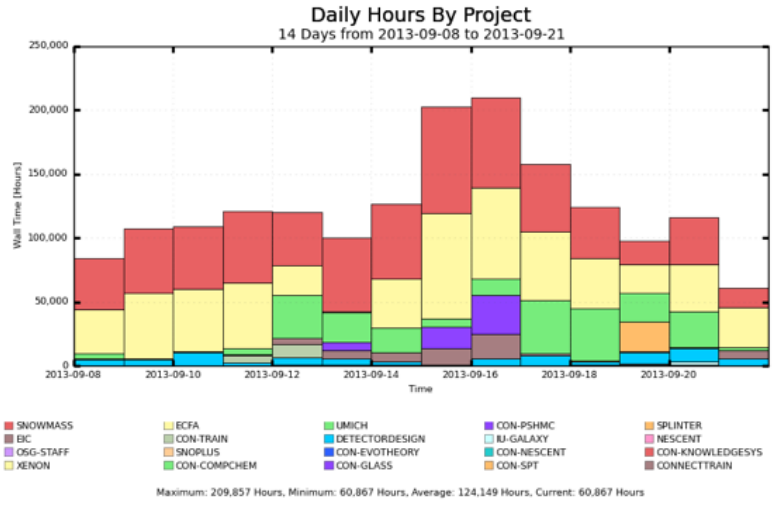

Figure 5: Daily resource usage of the OSG Resources by Campus Grid projects.

6. Integration With other CIs accounting services

6.1. XSEDE Accounting Service (AMIE API).

The OSG has been an XSEDE Service Provider since April, 2012. OSG provides a Virtual Cluster that shields XSEDE users from the necessity to be aware of the OSG infrastructure. Jobs submitted into

this Virtual Cluster are executed on the OSG physical clusters. The account records collected in Gratia, summarized by allocation (projects) are reported to the XSEDE account management system (AMIE) by using AMIE API.

6.2. EGI Accounting Service (SSM/APEL API)
Accounting records for US-LHC Tier1 and Tier2 facilities collected by Gratia are forwarded to the EGI accounting system (APEL), in accordance with signed MOU agreements. An external service parses and analyses accounting records for a resource, scales wall time and cpu usage, and then forwards this data to the APEL Server using the EGI GOC.

7.Future Development

7.1. Expansion of Collected Metrics.
Collected data is focused on CPU Usage. We are considering means to collect memory usage, as well as WAN and LAN network activities [6]. Currently for cloud accounting, the Gratia probe only interfaces with Open Nebula. We will be working on development of a generic interface to various other Cloud Management Services such as Open Stack, Eucalyptus, and others. We should be able to incorporate not only the static attributes and metrics of a VM but also dynamic usage of the resources such as cpu, memory, disk usage and networking activities into the accounting records. We are planning to account for persistent storage used by a VM as well as elastic storage used only during the VM lifetime.

7.2. Accounting & MultiGrids Infrastructure:
For the time being, means for information exchanged between various accounting services are limited. We are planning to explore the possibility to define a common interface that if accepted by various accounting services would allow Gratia to pull information across different distributed computing infrastructures and present a unified view of resources utilization.

Acknowledgments
The authors would like to thank Philippe Canal (Gratia architect /developer until 2011), John Weigand (OSG/EGI reporting), Mats Rynge (OSG/XD reporting), Soichi Hayashi (Project/OIM registration), and Brad Hurst (GratiaWeb development) for providing valuable support and development of the various parts of Gratia framework.

Fermilab is operated by the Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the United States Department of Energy.

References
[1] Canal P 2010 *Gratia: New Challenges in Grid Accounting J of Phys. Conference Series (JPCS)* **331** 1325
[2] Chadwick K et al. 2008 *FermiGrid — Experience and future plans J of Phys. Conference Series (JPCS)* **119 5** 052010
[3] Chadwick K, Timm S, Garzoglio G, Noh S 2013 *Grids, Virtualization and Clouds at Fermilab in these proceedings*
[4] Weitzel D 2011 *Campus Grid A Framework to Facilitate Resource Sharing Master thesis, University of Nebraska – Lincoln*
[5] Wilkins-Diehr N, Gannon D, Klimeck G, Oster S and Pamidighantam S 2008 *TeraGrid Science Gateways and Their Impact on Science. IEEE Computer* **41(11)** 32-41
[6] Bockelman B 2013 *Improved Site Accounting http://osg- docdb.opensciencegrid.org/cgi-bin/ShowDocument?docid=1083*